Summary: Assessment of Race, Ethnicity, Language and Disability (REALD) Data Quality in the Oregon Health Plan ONE System

Oregon Health
Authority
EQUITY AND INCLUSION DIVISION

Acknowledgments

How to cite this publication

Please cite this publication as follows:

McGee, MG. (2018). Summary: Assessment of race, ethnicity, language and disability (REALD) data quality in the Oregon Health Plan ONE system. Portland, Oregon: Oregon Health Authority, Equity and Inclusion Division.

This document can be provided upon request in an alternate format for individuals with disabilities or in a language other than English for people with limited English skills. To request this publication in another format or language, contact the Office of Equity and Inclusion at 971-673-3390 (voice or VRS), or email OHAREALD.Questions@odhsoha.oregon.gov.

Contents

Introduction

The Oregon Health Authority Equity and Inclusion (OEI) Division used one year of Oregon Health Plan applicant data from the Oregon Eligibility (ONE) system to conduct a detailed data quality analysis. The analysis addressed six overarching data quality questions. It resulted in 16 recommendations that will inform the Integrated Eligibility system's development. The analysis will also help develop and incorporate REALD into other data systems.

Background and purpose

In 2013, the Oregon Legislature passed [HB 2134](#). The law directed the Oregon Health Authority (OHA), collaborating with the Department of Human Services (DHS), to standardize and improve methods of collecting race, ethnicity, spoken and written language, and disability (REALD) demographics. This effort came from a need to address inconsistent and insufficient DHS and OHA data collection standards for these variables. Uniform and effective collection allows better measurement and comparison of disparities in services and health. Ultimately, this improves service delivery quality. See Appendix A for an at-a-glance view of these standards (in a paper-based survey format).

The data quality assessment of REALD data in ONE

Non-responses and other data quality issues regarding REALD questions emerged following REALD implementation. OEI conducted a data quality assessment to learn more about the data quality issues. Our goal was to avoid duplicating these issues in the IE system.

Methods

Evaluation questions

OEI sought to address the following six broad evaluation questions based on the REALD data collected from the applicant portal of new OHP members enrolled between Sept. 1, 2017, and June 30, 2018:

1. **Response rate overall:** To what extent are OHP enrollees answering the REALD questions?

2. **Race/ethnicity:** What can we infer from this analytic sample about enrollees who answered the race/ethnicity questions?

3. **Disability:** What can we infer from this analytic sample about enrollees who answered the disability questions?

4. **Language:** What can we infer from this analytic sample about enrollees who answered any of the language questions?

5. **Concordance (consistency):** To what extent are responses to the REALD questions consistent with each other?

6. **Limitations:** To what extent do the limitations that emerge from this assessment affect our ability to identify, measure and address inequities or disparities?

Sample

REALD data collection standards are not yet implemented across all MMIS data sources. As a result, OEI sampled enrollees new to the Maintenance Management Information System (MMIS) who enrolled between Sept. 1, 2017, and June 30, 2018. This produced an initial sample of 198,318 individuals. After excluding those not likely exposed to the REALD questions, 157,566 individuals (80% of the initial sample) made up the final analytic sample.

Key findings

The profile of MMIS enrollees does not reflect Oregonians' profile from the American Community Survey (ACS) data. This incongruity particularly holds for people with disabilities and people with limited English proficiency. Thus, we cannot use the MMIS data to generalize beyond the subset of enrollees. The REALD data's validity in MMIS for language and disability questions is inadequate because of the questions' format and design (i.e., being able to be skipped) rather than because of the questions themselves. The key system and process limitations follow.

Limitations stemming from the ONE system data collection design

Many of the data quality issues identified likely stem from the many "did not answer" responses. This resulted from respondents being able to skip REALD questions. Of the 157,566 individuals exposed to the REALD questions in the ONE system applicant portal, 33% answered all REALD key questions (which could include a decline or unknown response). More than one-third (34%) of the applicants answered the questions about racial/ethnic identity but none of the disability questions. Twenty-two percent answered the disability questions but not the questions about racial/ethnic identity. Five percent of enrollees in the analytic sample answered just the question about English proficiency and/or interpreter needs.

Limitations from lack of validation protocols

Certain patterns from enrollees' language access needs and disability responses make it hard to identify and address disparities. For example, it is hard to know what to infer based on non-congruent responses such as the following:

- Preferred spoken communications in English **and** needed an interpreter **and** does not need sign language interpretation.

- Needed both a language interpreter **and** sign interpreter.

- Speaks English "not well" or "not at all" **and** does not need interpretation.

- Needed written materials in alternate formats **and** does not have a disability.

- Needed sign interpretation **and** is not deaf or has serious difficulty hearing.

- Has an activity limitation **and** answered no to the other disability question.

- Has difficulty with dressing and bathing **and** doing errands on their own (age 15 and older) **and** does not have a limitation in any of four major domains (hearing, vision, cognitive and mobility).

We do not have validation protocols to ensure enrollees confirm responses that seem contradictory. As a result, we cannot make hypotheses or inferences to help understand the population being served. Further, while responses to the questions about preferred spoken and written languages and interpreter needs could be extremely helpful to assure language access (if validated), they do not lend themselves to being used in health disparity research.

System design and protocols limitations

Primary race. A major limitation relates to primary race. If a person with more than one racial or ethnic identity did not answer the primary race question, we do not use that individual's responses about their racial and ethnic identities. Because the primary race rule affects just those with more than one identity, these rules disproportionately affect the representation of people of color. Without using the race/ethnic identity data enrollees did provide, we lose the opportunity to improve the completeness of the data in analyses and reporting. To address this limitation, and for the purposes of this assessment, OEI re-created the primary race field. We used the "most identify/rarest" group methodology (1); this imputes a primary race if the person did not indicate one when reporting two or more races. For example, those who identified both as Western European and African American whose primary race field did not contain a racial/ethnic identity, their primary race identity would be assigned to the African American subgroup, based on which group was smaller in size (in Oregon).* These modifications resulted in recovering primary race data for 12,373 individuals. This decreased the number of "other race alone" by 578% and decreased the number of unknown primary race by 253%. The increase in aggregated racial/ethnic categories ranged from 38% (Latino/a and American Indian and Alaska Native) to 2% for Whites. We then compared how individuals answered racial and ethnic identity to those reporting receiving Medicaid (using ACS 2012–2016 estimates). The resulting aggregate figures are comparable to the ACS 2012–2016 figures for Oregon, using the same "most identify/rarest" group methodology.

Tracking data collection. We can improve data quality with targeted technical assistance and training if we know more about where most data quality issues occur. We can more fully rely on the data to inform our knowledge about disparities.

---

* This algorithm reflects populations at the state level and needs to be adjusted whenever focusing on a specific area of the state. These differences might lead to slight differences in identifying disparities.

However, the current system does not include data fields that would help address data quality by identifying the source of most of the issues. The lack of these data fields affects OHA and DHS staff members' ability to address data quality issues.

MMIS limitations handling REALD data from ONE

Some data quality issues result from how data flow from the ONE system to MMIS and then to Decision Support Surveillance and Utilization Review System (DSSURS). To fit or comply with the parameters of MMIS, some of the REALD data are changed or not captured at all. This process undermines the quality of the REALD data and the data's usefulness in identifying and addressing disparities. Below we identified four issues resulting from this process.

English as the default. The default to English when the applicant skipped the preferred written and/or spoken language means we cannot assume the person actually preferred English; this limits our ability to identify and address disparities associated with language barriers, access and discrimination. OHA staff also cannot ensure language access based on these responses. We learned that this default process results from Centers for Medicare & Medicaid Services (CMS)/MMIS requirements that do not allow for unknown responses in the preferred language questions.

Open-text fields. MMIS does not capture responses to questions in which the applicant writes in an open-text box, such as one of the race/ethnicity questions ("How do you identify your race, ethnicity, tribal affiliation, country of origin, or ancestry"). This means our learning is limited to emerging subgroups of racial/ethnic identities not currently used. We cannot assess congruence between responses to the open-ended question and the other race/ethnicity questions. We are also unable to discern the reason why people selected "other" racial/ethnic identity. We missed an opportunity to fill in applicable race and ethnic identity if the person answered the open-ended question but did not answer the other questions about racial and ethnic identities. These limitations weakened our ability to use the REALD data to identify and address disparities.

Time and date fields. Time and date fields for collecting demographic data from the enrollee do not carry over into MMIS. Therefore, we cannot track over time changes in how people identify their race and ethnicity, their ability to speak English, changes in needs for language access, and changes in limitations.

Prioritizing data fields from other systems. To complicate this further, the other data systems feeding into MMIS affect the data quality of the REALD responses. These systems do not send or populate the same REALD data elements. Updates triggered by the MMIS system may overwrite or leave different data elements in place. For example, instead of using the granular racial/ethnic identities reported in the applicant portal, we end up using older (usually aggregated) race information inconsistent with the REALD standards.

Recommendations

Recommendations

The assessment revealed several data quality issues organized in five categories. The recommendations listed below address these issues.

Data collection and data collection design recommendations

1. Require individuals to pick a response option (including decline or unknown response).

2. Use the "did not answer" response only for paper applications in which the applicant did not answer the question.

3. Combine the two interpreter questions into one. When one indicates a need for interpretation, have a follow-up question capture the type of interpreter or language.

4. Review the placement of all REALD questions, and the age follow-up question "did not answer." Have people with low vision check the visual placement of these questions on the screen, and the accessibility of these questions using screen readers.

Validation of response protocols

5. Language — Ask applicants to confirm their responses if the applicant(s):

   a. Prefer spoken communications in English and needed a language interpreter (the validation may need to be modified if the two interpreter questions are combined into one question if recommendation #3 is followed) and,

   b. State they do not speak English "well" or "very well" and they say they do not need interpretation. Let them know there is no charge for us to provide interpretation to speak with them in their preferred language.

6. Disability — Ask applicants to confirm their responses if they:

   a. State they need alternate formats but answer "no" to all disability questions

b. State they need sign interpretation but answer "no" to questions about being deaf or having serious difficulty hearing

c. State they have an activity limitation but answer "no" to the other disability question

d. State their current age is less than the age they acquired the limitation (for each disability question). Ask the applicant to select an age acquired that is equal to or less than current age and,

e. State they have difficulty with dressing and bathing and doing errands on their own (age 15 and older) but answer "no" to questions about hearing, vision, cognitive and mobility limitations.

System protocols

7. Remove the default of "English" in MMIS if the enrollee did not answer the questions about preferred written and/or spoken language. If this is not possible, see recommendation #14.

8. Include open-text fields in MMIS to analyze the open questions related to race/ethnicity. If this is not possible, see recommendation #14.

9. Develop internal data system protocols whereby legacy systems do not overwrite REALD information from the applicant portal. If this is not possible, see recommendation #14.

10. Add a date field for each set of REALD responses to track response changes over time. Important data is associated with changes over time related to racial and ethnic identities, language preferred, interpreter needs, English proficiency and functional limitations (disabilities). The date field with most current REALD information that OHP members provide gives analysts longitudinal data they can use to better understand members in the context of the social determinants of health, as well as identifies and addresses disparities in services and outcomes. If this is not possible, see recommendation #14.

11. Create separate fields for each racial/ethnic identity. If this is not possible, see recommendation #14.

12. Build in a mechanism to periodically track and assess data quality by:

a. How the data were collected (e.g., applicant portal, worker portal with case worker or assistor, paper application), and

b. By whom (e.g., the applicant applied directly through the online applicant portal vs a case worker or community assistor helping submit the application).

In addition, capture the case worker's and community assistor's location and identification to target data quality interventions.

13. Revise primary race "rules" to use "most identify/rarest" group methodology if there were two or more racial/ethnic identities and the person did not select a primary race field.

Alternate data systems and processes to retrieve and store REALD data

14. Because of the system's limitations noted in these recommendations, consider if MMIS should be the main source of data for REALD. A separate REALD database with access to the date/time fields from ONE and the original REALD data before being changed by MMIS may be easier and more cost effective than trying to make REALD fit into MMIS in ways MMIS is not able to handle. If this is possible, it may be a better way of providing quality REALD data to the CCOs and OHA researchers to meet the goal of HB 2134: to identify and address disparities.

Continuous quality improvement processes and exploratory research

15. Establish a continuous quality improvement (CQI) focusing on the quality of the REALD data in ONE and later in IE. REALD subject matter experts from OEI and OEMS should be fully involved. It is essential those who can authorize and ensure implementation of the CQI team's recommendations are also involved and support the process. Specifically, this team could:

    a. Assess the impact on REALD data quality each time improvements are made to the ONE/IE system within three to six months

    b. Engage with recommendations in this report and,

    c. Learn more about what and why some of the data quality issues are occurring, implement solutions and assess the solution(s)' effectiveness.

16. Conduct exploratory research based on some of the issues raised in this report (e.g., recommendation #4). The CQI team should lead this effort. Specifically, the CQI team should:

    a. Learn more about needs for alternate formats, particularly among people who answered "no" to all disability questions, and those with limited English proficiency

b. Learn about needs for additional granular racial-ethnic categories based on responses to the open-ended question about racial and ethnic identity, particularly for those who chose an "other" subgroup category (e.g., Other Pacific Islander, Other Asian, Other White) and,

c. Conduct participant observations and focus groups with community assistors and those working in the call centers to learn how they are interacting with applicants regarding REALD questions.

Immediate opportunities for improvement

The assessment revealed a need to make changes to the ONE/IE system, along with the 16 recommendations listed above. We can make some changes to the IE system before it launches that could dramatically improve both data quality and the data's usefulness in identifying and addressing disparities.

Five of the 16 recommendations should be implemented sooner than later, and preferably before the launch of the IE system in 2020. Doing so should greatly improve data quality with a minimum of costs and effort. These five recommendations are:

1. Require individuals to pick a response option (including "decline" or "unknown response").

2. Have the "did not answer" response only available as an option when the paper application is reviewed manually, and it is clear that the applicant did not answer the question.

3. Remove the default of "English" in the MMIS system if the enrollee did not answer the questions about preferred written and/or spoken language.

4. Because of the system's limitations noted in other recommendations, consider if MMIS should be the main source of data for REALD. A separate REALD database with access to the date/time fields from ONE and the original REALD data before being changed by MMIS may be easier and more cost effective than trying to make REALD fit into MMIS in ways MMIS is not able to handle.

5. Establish a continuous quality improvement (CQI) team focusing on REALD data quality in ONE and later in IE.

Endnotes

1. Mays VM, Ponce NA, Washington DL, Cochran SD. Classification of race and ethnicity: implications for public health. Annual Review of Public Health. 2003;24: 84–110.

Appendix A. At-a-glance view of REALD questions and categories

Race, Ethnicity, Language, and Disability (REALD)

Oregon Health Authority

These questions are optional and your answers are confidential. We would like you to tell us your race, ethnicity, language and disability background so that we can find and address health and service differences.

1. Do you need written materials in an alternate format (Braille, large print, audio recordings, etc.)?

☐ Yes ☐ No ☐ Don't know/Unknown ☐ Don't want to answer/Decline

If yes, which format? _____

Race and Ethnicity

2. How do you identify your **race, ethnicity, tribal affiliation, country of origin, or ancestry?**

_____

3. Which of the following describes your **racial or ethnic identity?** Please check **ALL** that apply.

American Indian or Alaska Native
- ☐ American Indian
- ☐ Alaska Native
- ☐ Canadian Inuit, Metis, or First Nation
- ☐ Indigenous Mexican, Central American, or South American

Hispanic or Latino/a
- ☐ Hispanic or Latino/a Central American
- ☐ Hispanic or Latino/a Mexican
- ☐ Hispanic or Latino/a South American
- ☐ Other Hispanic or Latino/a

Asian
- ☐ Asian Indian
- ☐ Chinese
- ☐ Filipino/a
- ☐ Hmong
- ☐ Japanese
- ☐ Korean
- ☐ Laotian
- ☐ South Asian
- ☐ Vietnamese
- ☐ Other Asian

Native Hawaiian or Pacific Islander
- ☐ Guamanian or Chamorro
- ☐ Micronesian*
- ☐ Native Hawaiian
- ☐ Samoan
- ☐ Tongan*
- ☐ Other Pacific Islander

Black or African American
- ☐ African American
- ☐ African (Black)
- ☐ Caribbean (Black)
- ☐ Other Black

Middle Eastern/Northern African
- ☐ Northern African
- ☐ Middle Eastern

White
- ☐ Eastern European
- ☐ Slavic
- ☐ Western European
- ☐ Other White

Other Categories
- ☐ Other (please list)

  _____
- ☐ Don't know/Unknown
- ☐ Don't want to answer/Decline

4. If you selected more than one racial or ethnic identity above, please **CIRCLE the ONE that best represents your racial or ethnic identity.** If you have more than one primary racial or ethnic identity please check here: ☐

You can get this document in other languages, large print, braille, or a format you prefer. We accept all relay calls or you can dial 711. Contact:

Program:

Phone:

Email:

OHA 0074 (9/18)

Language

5. In what **language** do you want us to:

   Speak with you _____

   Write to you _____

6. Do you need a **sign language** interpreter for us to communicate with you?

   ☐ Yes   ☐ Don't know/Unknown
   ☐ No   ☐ Don't want to answer/Decline

   If yes, which type do you need us to communicate with you?
   (ASL, PSE, tactile interpreting, etc.)

   _____

7. Do you need an **interpreter** for us to communicate with you?

   ☐ Yes   ☐ Don't know/Unknown
   ☐ No   ☐ Don't want to answer/Decline

8. How well do you speak English?

   ☐ Very Well   ☐ Not at all
   ☐ Well   ☐ Don't know/Unknown
   ☐ Not Well   ☐ Don't want to answer/Decline

Disability   Your answers will help us find health and service differences among people with and without functional difficulties. Your answers are confidential.

9. Are you **deaf** or do you have **serious difficulty hearing?**

   ☐ Yes   ☐ Don't know/Unknown
   ☐ No   ☐ Don't want to answer/Decline
   If yes, at what age did this condition begin? _____

10. Are you **blind** or do you have **serious difficulty seeing,** even when wearing glasses?

    ☐ Yes   ☐ Don't know/Unknown
    ☐ No   ☐ Don't want to answer/Decline
    If yes, at what age did this condition begin? _____

11. Does a **physical, mental, or emotional condition limit your activities** in any way?

    ☐ Yes   ☐ Don't know/Unknown
    ☐ No   ☐ Don't want to answer/Decline
    If yes, at what age did this condition begin? _____

12. What is your age today? _____

    Please stop now if the person is under age 5

13. Do you have serious difficulty **walking or climbing stairs?**

    ☐ Yes   ☐ Don't know/Unknown
    ☐ No   ☐ Don't want to answer/Decline
    If yes, at what age did this condition begin? _____

14. Do you have **difficulty dressing or bathing?**

    ☐ Yes   ☐ Don't know/Unknown
    ☐ No   ☐ Don't want to answer/Decline
    If yes, at what age did this condition begin? _____

15. Because of a **physical, mental, or emotional condition,** do you have serious difficulty:

    a. **Concentrating, remembering or making decisions?**

    ☐ Yes   ☐ Don't know/Unknown
    ☐ No   ☐ Don't want to answer/Decline
    If yes, at what age did this condition begin? _____

    Please stop now if you/the person is under age 15

    b. **Doing errands alone** such as visiting a doctor's office or shopping?

    ☐ Yes   ☐ Don't know/Unknown
    ☐ No   ☐ Don't want to answer/Decline
    If yes, at what age did this condition begin? _____

OHA 0074 (9/18)